Computationally simple MMSE (A-optimal) Adaptive Beam-pattern Design for MIMO Active Sensing Systems via a Linear-Gaussian Approximation

Computationally simple MMSE (A-optimal) adaptive beam-pattern design for MIMO active sensing systems via a linear-Gaussian approximation

Steven Herbert, James R. Hopgood, *Member, IEEE,* Bernard Mulgrew, *Fellow, IEEE*

***Abstract*—This paper presents an approximate minimum mean squared error (MMSE) adaptive beam-pattern design (ABD) method for MIMO active sensing systems. The proposed approximate MMSE ABD method leverages the physics of the MIMO arrays to provide a linear-Gaussian approximation that is specific to MIMO active sensing systems, and yields a computationally simple optimisation problem. Computational complexity analysis confirms this theoretical reduction in the number of floating-point operations required, most notably that evaluation of the proposed approximate optimisation cost function grows polynomially with the number of targets being tracked, whereas for evaluation of the exact cost the growth is exponential. Additionally, numerical results indicate that, even for a simple scenario with a single target being tracked, the proposed approximate MMSE ABD method does indeed reduce the mean squared error of target parameter estimation compared to the non-adaptive case, with a reduction in computation time of four orders of magnitude compared to exact MMSE ABD.**

***Index Terms*—Adaptive waveform design, adaptive beam-pattern design, adaptive beamforming, minimum mean squared error, active sensing, MIMO, radar, Bayesian, particle filters, optimal design, adaptive beam-forming.**

I. Introduction

ADAPTIVE beam-pattern design (ABD) in active sensing systems is a currently active area of research. In particular, there is interest in the development of cognitive radar systems [1]–[3], in which one important feature is the use of the current estimate of target parameters to design the next transmitted beam-pattern such that it is expected to improve the target parameter estimation after the next measurement (received reflected signal). Huleihel *et al* show the general architecture of ABD in active sensing systems [4, Fig. 1]. In this paper, we are concerned with adaptive shaping the beam-pattern generated by a multiple-input-multiple-output (MIMO) array, which is also known as adaptive waveform design in some of the iterature [4], [5]. Specific MIMO active sensing system modalities to which this analysis applies include MIMO sonar [6] as well as the aforementioned MIMO radar. A comparison of the various common optimisation criteria for problems of this sort led to the conclusion that minimising the expected mean squared error is a good overall choice of cost function [7], and hence it is this criteria that we employ in this work. In signal processing this is generally denoted minimum mean squared error (MMSE) design, whereas in the optimal design literature this is referred to as 'A-optimal' design [8].

The cost function for MMSE design in active sensing systems has been expressed [5, Eq. (11)], however evaluation thereof is shown to be computationally expensive and thus in this paper we provide an approximate method for MMSE ABD that is less computationally expensive than the exact solution. Our approach is to use the current estimates of the target parameters to artificially construct a multivariate Gaussian distribution for the unique terms of the channel response (a non-linear function of the target parameters), and to optimally design the beam-pattern for the resultant linear Gaussian system accordingly. This approximate channel model is expressed in the standard form for linear Gaussian channels [9], and indeed yields an optimisation problem that is shown to be significantly less computationally complex than the exact MMSE ABD algorithm [5], and also the approximate MMSE ABD algorithm proposed by Huleihel *et al* [4], which has similar complexity to the exact MMSE ABD algorithm [5, Fig. 7]. Notably, the ABD method proposed here explicitly uses the physics of the MIMO array to construct the Gaussian covariance matrix, which distinguishes it from techniques for handling non linear-Gaussian problems such as the extended Kalman filter [10] and Unscented Kalman filter [11], [12].

Additionally, it is pertinent that the approximation detailed herein requires that the target parameters are fully estimated. That is both the angle at which it is located and its complex attenuation – in the numerical results presented in [5] only the location angle is estimated. Thus, an additional significant contribution of this work is that the results included herein represent a shift towards an actual implementable system, in so far as that we consider a more representative physical model.

A. Contributions

The main contributions of this work are:

- We propose an approximate MMSE ABD method for MIMO active sensing systems.
- We provide numerical results that demonstate that our proposed method improves the target parameter estimation relative to the non-adaptive case, for a much reduced computational cost compared to the exact MMSE ABD method.
- We include analysis of the computational complexity to confirm that this computational saving is an essential

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Grant number EP/K014277/1 and the MOD University Defence Research Collaboration (UDRC) in Signal Processing.

S. Herbert, J. R. Hopgood and B. Mulgrew are with the Institute for Digital Communications, School of Engineering, University of Edinburgh, Edinburgh, UK (e-mail: S.Herbert@ed.ac.uk, James.Hopgood@ed.ac.uk, Bernie.Mulgrew@ed.ac.uk).

property of the algorithm and will therefore be present in any implementation. Notably, for the exact method [5], the number of operations required for cost function evaluation grows exponentially with the number of targets being tracked, whereas for the approximate method proposed herein evaluation of the cost function grows polynomially with the number of targets.
- We take an important step towards a MMSE ABD algorithm that can be deployed in an actual MIMO active sensing system by estimating the target attenuation as well as angle.

B. Notation

Our aim is to use standard and simple notation in this paper. In some parts of the analysis it is convenient to express functions and variables as some letter 'primed'. This is used to denote a variation of the primed variable/function and does not denote either differentiation, which is always explicitly expressed, or the Hermitian transpose, which we express $(\cdot)^H$. Other notation used in this paper includes transpose, which we express $(\cdot)^T$ and complex conjugation, which we express $(\cdot)^*$. The Kronecker product is denoted $\otimes$, and on occasion it is necessary to define a single (ith) element of a sum/ set as $(.)^{(i)}$, which should not be mistaken for a raised power, where the superscript is not in parentheses. Finally, note that $\mathbf{a}$ and $\boldsymbol{a}$ are not the same.

C. Paper organisation

The remainder of the paper is organised as follows: in Section II we define the MIMO active sensing system model; in Section III we derive and express the linear-Gaussian approximation of the MMSE cost function; in Section IV we show how this cost function can be optimised; in Section V we present our main results; in Section VI we present and discuss the computational complexity of the approximate MMSE ABD method we propose herein relative to the exact MMSE ABD method [5]; and finally in Section VII we draw conclusions.

II. System Model

We use the same system model as that considered when expressing the exact MMSE cost function [5], itself based on that used by Huleihel *et al* [4]. Our expressions concern the kth step (opportunity to adaptively design the beam-pattern), which consists of L different beam-patterns (snapshots). By definition, the MIMO active sensing system consists of N_T transmit elements and N_R co-located receive elements. $\mathbf{S}_k \in \mathbb{C}^{N_T \times L}$ denotes the transmitted beam-pattern, the lth column of which is a column vector corresponding to the lth snapshot (where $l = 1, \ldots, L$), and whose rows correspond to the complex signal transmitted at the given snapshot by each of the N_T transmitting elements. $\mathbf{X}_k \in \mathbb{C}^{N_R \times L}$ denotes the received beam-pattern, and again the columns correspond to the snapshots, with each row corresponding to the complex signal received on the respective receiving element. It follows that the channel is defined by:

$$\mathbf{X}_k = \mathbf{H}_k(\boldsymbol{\theta}_k)\mathbf{S}_k + \mathbf{N}_k, \tag{1}$$

where $\mathbf{H}_k(\boldsymbol{\theta}_k) \in \mathbb{C}^{N_R \times N_T}$ represents the channel response as a non-linear function (in general) of $\boldsymbol{\theta}$, a vector of the Q parameters of the target, i.e., $\boldsymbol{\theta}_k \in \mathbb{C}^{Q \times 1}$. For simplicity, we assume that the target parameters do not vary within any given step. $\mathbf{N}_k \in \mathbb{C}^{N_R \times L}$ denotes additive white Gaussian noise (AWGN). The noise is circularly symmetric complex, i.e., each element of $\mathbf{N}_k$ is a complex number whose real part is an independent zero mean Gaussian random variable with variance σ_n^2 and whose imaginary part is also an independent zero mean Gaussian random variable with variance σ_n^2, and the various elements of $\mathbf{N}_k$ are mutually independent. This channel model represents the situation where the received signal is a linear function of the transmitted signal (plus AWGN), but a non-linear function of the model parameters.

In general, the target parameters may vary from step to step, according to a statistical process:

$$\boldsymbol{\theta}_k = \mathbf{f}(\boldsymbol{\theta}_{k-1}, \mathbf{v}_{k-1}), \tag{2}$$

where $\mathbf{f}(.)$ is an arbitrary function and $\mathbf{v}_{k-1}$ is noise, which is independent of $\mathbf{N}_k$. The formulation developed in this paper would apply if $\mathbf{f}(.)$ were to change at each step, however to simplify the notation in the following analysis we fix $\mathbf{f}(.)$. For simplicity, we do not allow a mismatch between the actual target motion and the statistical model available to the MIMO active sensing system. It is, however, worth noting that the results in [13] indicating that MMSE ABD is still effective even when there is such a mismatch.

On a similar note, a more physically reasonable model would include the possibility of signal dependent interference, however we do not include this here for consistency with the formulations used in [4], [5], [7], [13]. It should, however, be noted that there exists literature which does address this scenario [14]–[17]. Likewise, we assume that calibration, range and Doppler measurements are achieved using conventional methods [18], and their specific realisation is independent of the ABD method proposed herein and therefore outside of the scope of this paper.

III. Expression of a linear-Gaussian approximate cost function for ABD

As related in Section I, the cost function for MMSE ABD can be expressed exactly [5, Eq. (11)], however numerical evaluation thereof is computationally expensive. Specifically, this high computational cost arises because of the existence of a double integral that is approximately evaluated using two nested sums over a large number of particles / samples in the implementation. Clearly there is a benefit in finding a computationally simple approximate alternative, and to achieve this it is first necessary to take a fresh look at the basics of the non-linear parameter estimation.

A. Determination of $\boldsymbol{\theta}_k$ from $\mathbf{H}(\boldsymbol{\theta}_k)$

First, we address the capability of the MIMO active sensing system to determine $\boldsymbol{\theta}_k$ from $\mathbf{H}(\boldsymbol{\theta}_k)$, by addressing how the physics of the MIMO arrays enables the calculation of $\boldsymbol{\theta}_k$.

From the standard form of MIMO active sensing systems [4], [5], [19]:

$$\mathbf{H}(\boldsymbol{\theta}_k) = \sum_{q=1}^{Q'} [\boldsymbol{\alpha}_k]_q \mathbf{a}_R([\boldsymbol{\phi}_k]_q)\mathbf{a}_T^T([\boldsymbol{\phi}_k]_q), \tag{3}$$

where $\boldsymbol{\theta}_k = [\boldsymbol{\phi}_k; \Re(\boldsymbol{\alpha}_k); \Im(\boldsymbol{\alpha}_k)]$ in which $\boldsymbol{\phi}_k \in \mathbb{R}^{Q'\times 1}$ is a vector of the target angles and $\boldsymbol{\alpha}_k \in \mathbb{C}^{Q'\times 1}$ is a vector of the target attenuations, where Q' is the number of targets, and $\mathbf{a}_T(\cdot)$ and $\mathbf{a}_R(\cdot)$ are the steering vectors of the transmit and receive arrays, respectively. For ease of exposition, let both the transmit and receive arrays have the same spacing, $\tilde{\lambda}$ (we later relate our analysis to the general case), then:

$$\mathbf{a}_R([\boldsymbol{\phi}_k]_q) = [[\boldsymbol{a}_k]_q^0, [\boldsymbol{a}_k]_q^1, \ldots, [\boldsymbol{a}_k]_q^{N_R-1}]^T, \tag{4}$$

$$\mathbf{a}_T([\boldsymbol{\phi}_k]_q) = [[\boldsymbol{a}_k]_q^0, [\boldsymbol{a}_k]_q^1, \ldots, [\boldsymbol{a}_k]_q^{N_T-1}]^T, \tag{5}$$

where $[\boldsymbol{a}_k]_q = \exp(2\pi\sqrt{-1}(\tilde{\lambda}/\lambda)\sin([\boldsymbol{\phi}_k]_q))$, in which λ is the wavelength. Thus the qth term of the sum in (3), defined as $\mathbf{H}^{(q)}$, can be expressed:

$$\begin{aligned}\mathbf{H}^{(q)} &= [\boldsymbol{\alpha}_k]_q \mathbf{a}_R([\boldsymbol{\phi}_k]_q)\mathbf{a}_T^T([\boldsymbol{\phi}_k]_q) \\ &= [\boldsymbol{\alpha}_k]_q \begin{bmatrix} [\boldsymbol{a}_k]_q^0 & [\boldsymbol{a}_k]_q^1 & \ldots & [\boldsymbol{a}_k]_q^{N_T-1} \\ [\boldsymbol{a}_k]_q^1 & [\boldsymbol{a}_k]_q^2 & \ldots & [\boldsymbol{a}_k]_q^{N_T} \\ \vdots & \vdots & \ddots & \\ [\boldsymbol{a}_k]_q^{N_R-1} & [\boldsymbol{a}_k]_q^{N_R} & & [\boldsymbol{a}_k]_q^{N_T+N_R-1} \end{bmatrix}\end{aligned} \tag{6}$$

for $q = 1, \ldots, Q'$. We now consider the determination of $[\boldsymbol{\phi}_k]_q$. Following convention [4], [5], we measure angle in degrees, and by definition $-90 \leq [\boldsymbol{\phi}_k]_q < 90$ thus $[\boldsymbol{\phi}_k]_q$ can only be uniquely determined from the complex value of $[\boldsymbol{a}_k]_q$ if $\tilde{\lambda} \leq \lambda/2$. So it follows that we must determine $2Q'$ complex numbers in order to find $\boldsymbol{\theta}_k$ (i.e., each of the Q' targets is parameterised in terms of $[\boldsymbol{a}_k]_q$ and $[\boldsymbol{\alpha}_k]_q$). According to (6), if $\mathbf{H}^{(q)}$ is exactly known for all q (i.e., $\mathbf{H}(\boldsymbol{\theta}_k)$ is exactly known), it is composed of $N_T + N_R - 1$ distinct complex numbers, and thus $\boldsymbol{\theta}_k$ can be determined if, and only if, $N_T + N_R - 1 \geq 2Q'$. However, the simultaneous equations are polynomial and thus there are potentially multiple solutions (roots) spread across the complex plane. We know that the relevant solution has the property that $|[\boldsymbol{a}_k]_q| = 1$ by definition, and we assume that this additional constraint is sufficient to identify which of the solutions is the 'correct' one, thus effectively meaning we have a procedure leading to a unique solution. This solution can be justified theoretically by the fact that the probability of an arbitrary root (i.e., other than the 'correct one') having the property $|[\boldsymbol{a}_k]_q| = 1 + \epsilon$ vanishes as the magnitude of ϵ tends to zero. This theoretical justification is based on the is supported by the empirical evidence that MIMO active sensing systems *do* enable the estimation of target parameters, as exemplified by the results included in Section V, and indeed in the abundance of literature on the subject (including [4]–[7] etc.).

It is important to note that this is a purely algebraic argument for the determination of $\boldsymbol{\theta}_k$, assuming that $\mathbf{H}(\boldsymbol{\theta}_k)$ is known. As such, it does not depend on $\mathbf{S}_k$ or L, which can be varied to improve the estimation of $\mathbf{H}^{(q)}$ through noisy measurement, but cannot help to resolve $\boldsymbol{\theta}_k$ if the underlying system is under-determined (i.e., if there are too many targets for the size of the MIMO arrays).

B. Re-arrangement of cost function in terms of $\boldsymbol{\psi}$

Let $\boldsymbol{\psi}_k \in \mathbb{C}^{(N_R+N_T-1)\times 1}$ be such that

$$[\boldsymbol{\psi}_k]_i = \sum_{q=1}^{Q'} [\boldsymbol{\psi}_k^{(q)}]_i, \tag{7}$$

where

$$[\boldsymbol{\psi}_k^{(q)}]_i = [\boldsymbol{\alpha}_k]_q [\boldsymbol{a}_k]_q^i \tag{8}$$

i.e., $[\psi_k]_i$ is a vector of the unique elements of $\mathbf{H}(\boldsymbol{\theta}_k)$. Accordingly, we can alternatively express the state-space definition of the channel,

$$\mathbf{X}_k = \mathbf{H}(\boldsymbol{\theta}_k)\mathbf{S}_k + \mathbf{N}_k, \tag{9}$$

as

$$\mathrm{vec}(\mathbf{X}_k) = \mathbf{S}'_k \boldsymbol{\psi}_k + \mathrm{vec}(\mathbf{N}_k), \tag{10}$$

where $\mathbf{S}'_k = [\mathbf{S}''_1; \mathbf{S}''_2; \ldots; \mathbf{S}''_L]$, in which $\mathbf{S}''_l \in \mathbb{C}^{N_R\times(N_T+N_R-1)}$ (where $1 \leq l \leq L$) is such that its jth row has the form $[\mathbf{0}_{1,j-1}, [\mathbf{S}_k]_{1l}, [\mathbf{S}_k]_{2l}, \ldots, [\mathbf{S}_k]_{N_T l}, \mathbf{0}_{1,N_R-j}]$, and $\mathbf{0}_{1,j'}$ is a row vector of zeros of length j'. This re-arrangement is a standard property of the *Hankel* matrix [20], of which (6) is an instance. For example, let $N_T = N_R = 3$ and $L = 1$, we have that:

$$\begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} \psi_1 & \psi_2 & \psi_3 \\ \psi_2 & \psi_3 & \psi_4 \\ \psi_3 & \psi_4 & \psi_5 \end{bmatrix} \begin{bmatrix} s_{11} \\ s_{21} \\ s_{31} \end{bmatrix} + \begin{bmatrix} n_{11} \\ n_{21} \\ n_{31} \end{bmatrix}, \tag{11}$$

which we can express:

$$\begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} s_{11} & s_{21} & s_{31} & 0 & 0 \\ 0 & s_{11} & s_{21} & s_{31} & 0 \\ 0 & 0 & s_{11} & s_{21} & s_{31} \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \\ \psi_5 \end{bmatrix} + \begin{bmatrix} n_{11} \\ n_{21} \\ n_{31} \end{bmatrix}. \tag{12}$$

Note that in each of (11) and (12) we have omitted the subscript k to aid readability.

Such a formulation again leads us to address the subject of the determination of $\boldsymbol{\theta}_k$, this time from $\mathbf{X}_k$. For observing (12), we can see that with $L = 1$ the linear system is under-determined, and thus it is not possible to determine $\boldsymbol{\psi}_k$ from $\mathbf{X}_k$. The subtlety here is that $\mathbf{S}'_k$ is not 'inverted' to estimate $\boldsymbol{\psi}_k$, but rather in a Bayesian framework $\mathbf{S}_k$ is designed to minimise the expected mean squared error after the next received reflected signal, $\mathbf{X}_{k+1}$. As we shall see, it is the correlations between the elements of $\boldsymbol{\psi}_k$ that enable us to do this. Importantly, however, if the beam-pattern is not adapted and remains constant throughout, then it *is* necessary to have $\mathbf{S}_k$ with a sufficient number of linearly independent columns. Typically this requirement is achieved by the sufficient condition that $L \geq N_R$, with mutually orthogonal columns.

$$\mathbf{R}_k^{(q)} = \text{var}([\boldsymbol{\alpha}_k]_q) \begin{bmatrix} 1 & \langle [\boldsymbol{a}_k]_q \rangle^* & \langle [\boldsymbol{a}_k]_q^2 \rangle^* & \dots & \langle [\boldsymbol{a}_k]_q^{N_T+N_R-2} \rangle^* \\ \langle [\boldsymbol{a}_k]_q \rangle & 1 & \langle [\boldsymbol{a}_k]_q \rangle^* & \dots & \langle [\boldsymbol{\alpha}_k]_q^{N_T+N_R-3} \rangle^* \\ \langle [\boldsymbol{\alpha}_k]_q^2 \rangle & \langle [\boldsymbol{\alpha}_k]_q \rangle & 1 & \dots & \langle [\boldsymbol{\alpha}_k]_q^{N_T+N_R-4} \rangle^* \\ \vdots & \vdots & \vdots & \ddots & \\ \langle [\boldsymbol{\alpha}_k]_q^{N_T+N_R-2} \rangle & \langle [\boldsymbol{\alpha}_k]_q^{N_T+N_R-3} \rangle & \langle [\boldsymbol{\alpha}_k]_q^{N_T+N_R-4} \rangle & & 1 \end{bmatrix} \tag{21}$$

C. Treating $\boldsymbol{\psi}_k$ as multivariate complex Gaussian

To manipulate $\boldsymbol{\psi}_k$ to develop a computationally efficient ABD algorithm, it is first necessary to define a suitable model for $\boldsymbol{\theta}_0$, that is the distribution for the parameters prior to any received signals. For consistency with the literature [4], [5], we define:

$$[\boldsymbol{\phi}_0]_q \sim \mathcal{U}(-90^o, 90^o), \tag{13}$$

where $q = 1, \dots, Q'$. That is, *a priori* we treat all targets to be uniformly distributed between -90^o and 90^o. In the results presented for the exact MMSE optimisation [5], $\boldsymbol{\alpha}$ was assumed to be known, however in the work by Huleihel *et al* the *a priori* distribution of the target attenuation was treated as a circularly symmetric complex Gaussian random variable for each target:

$$[\boldsymbol{\alpha}_0]_q \sim \mathcal{CN}(0, \text{var}([\boldsymbol{\alpha}_0]_q)), \tag{14}$$

where $q = 1, \dots, Q'$ and all targets angles and attenuations are independent. We also have that

$$|[\boldsymbol{a}_k]_q| = |\exp(2\pi\sqrt{-1}(\tilde{\lambda}/\lambda)\sin([\boldsymbol{\phi}_k]_q))| = 1, \tag{15}$$

i.e., a complex number with magnitude 1. So it follows that substituting (14) and (15) into (8) (i.e., for $k = 0$) and using the phase invariance of the circularly symmetric complex Gaussian distribution [21, Definition 3.7.2] yields:

$$[\boldsymbol{\psi}_0^{(q)}]_i = [\boldsymbol{a}_0]_q [\boldsymbol{\alpha}_0]_q \sim \mathcal{CN}(0, \text{var}([\boldsymbol{\alpha}_0]_q)), \tag{16}$$

and using the standard summing properties of independent Gaussian distributions to substitute (16) into (7) yields:

$$[\boldsymbol{\psi}_0]_i \sim \mathcal{CN}\left(0, \sigma_0^2\right), \tag{17}$$

where $\sigma_0^2 = \Sigma_{q=1}^{Q'} \text{var}([\boldsymbol{\alpha}_0]_q)$ for all i.

Whilst each element of $\boldsymbol{\psi}_0$ is distributed as a circularly symmetric complex Gaussian (as specified in (17)), it does not follow that the joint distribution of $\boldsymbol{\psi}_0$ is necessarily a multivariate complex Gaussian (e.g., [22, pp. 372–373]). However, in our approximate method we depart from the actual mathematical nature of $\boldsymbol{\psi}_0$ and assume that $\boldsymbol{\psi}_0$ *is* a multivariate circularly symmetric complex Gaussian:

$$\boldsymbol{\psi}_0 \sim \mathcal{CN}\left(\mathbf{0}, \sigma_0^2 \mathbf{I}_{N_T+N_R-1}\right). \tag{18}$$

Moreover, our method is to treat $\boldsymbol{\psi}_k$ as multivariate circularly symmetric Gaussian for all k, for we know that this will yield a relatively simple optimisation problem. That is,

$$\boldsymbol{\psi}_k \sim \mathcal{CN}\left(\boldsymbol{\mu}_k, \mathbf{R}_k\right), \tag{19}$$

where $\boldsymbol{\mu}_k$ is not required, and:

$$\mathbf{R}_k = \sum_{q=1}^{Q'} \mathbf{R}_k^{(q)}, \tag{20}$$

where $\mathbf{R}_k^{(q)}$ is defined in (21), in which $\langle [\boldsymbol{a}_k]_q^i \rangle \triangleq \mathbb{E}([\boldsymbol{a}_k]_q^i | \mathbf{X}^{k-1})$, determined by $p(\boldsymbol{\theta}_k | \mathbf{X}^{k-1})$. As in the exact MMSE ABD method [5], $p(\boldsymbol{\theta}_k | \mathbf{X}^{k-1})$ is available (or approximately available) from the underlying estimation of $\boldsymbol{\theta}_k$, for example by a particle filter (PF). Likewise, $\text{var}([\boldsymbol{\alpha}_k]_q)$ is found in the same manner. Specifically, we treat each element of $[\boldsymbol{\psi}_k^{(q)}]$, defined in (8), as a circularly symmetric complex Gaussian, where the mean is not required, and the variance used is that of the current PDF of the target parameters:

$$[\boldsymbol{\psi}_k^{(q)}]_i \sim \mathcal{CN}(\mu_i^{(q)}, \text{var}([\boldsymbol{\alpha}_k]_q)). \tag{22}$$

For each target the covariance matrix, $\mathbf{R}_k^{(q)}$ has been constructed from point estimates of $[\boldsymbol{a}_k]_q^i$. Formally, for qth target, we consider the covariance between the ith and jth elements (where $j > i$) and express the jth element as a sum of a correlated and uncorrelated component of the ith element:

$$[\boldsymbol{\psi}_k^{(q)}]_j = \langle [\boldsymbol{a}_k]_q^{j-i} \rangle [\boldsymbol{\psi}_k^{(q)}]_i + \tilde{\psi}, \tag{23}$$

where

$$\tilde{\psi} \sim \mathcal{CN}(\tilde{\mu}, \tilde{\sigma}^2) \tag{24}$$

from which we can express:

$$\begin{aligned}
[\mathbf{R}_k^{(q)}]_{i,j} &= \mathbb{E}\left(\left([\boldsymbol{\psi}_k^{(q)}]_i - \mu_i^{(q)} \right) \left([\boldsymbol{\psi}_k^{(q)}]_j - \mu_j^{(q)} \right)^* \right) \\
&= \mathbb{E}\Big(\left([\boldsymbol{\psi}_k^{(q)}]_i - \mu_i^{(q)} \right) \\
&\qquad \left(\langle [\boldsymbol{a}_k]_q^{j-i} \rangle [\boldsymbol{\psi}_k^{(q)}]_i + \tilde{\psi} - \langle [\boldsymbol{a}_k]_q^{j-i} \rangle \mu_i^{(q)} - \tilde{\mu} \right)^* \Big) \\
&= \left(\langle [\boldsymbol{a}_k]_q^{j-i} \rangle \right)^* \left(\mathbb{E}\left([\boldsymbol{\psi}_k^{(q)}]_i [\boldsymbol{\psi}_k^{(q)}]_i^* \right) - \mu_i^{(q)} \mu_i^{(q)*} \right) \\
&\quad + \left(\mathbb{E}\left([\boldsymbol{\psi}_k^{(q)}]_i \tilde{\psi}'^* \right) - \mu_i^{(q)} \tilde{\mu}^* \right) \\
&= \left(\langle [\boldsymbol{a}_k]_q^{j-i} \rangle \right)^* \left(\mathbb{E}\left([\boldsymbol{\psi}_k^{(q)}]_i [\boldsymbol{\psi}_k^{(q)}]_i^* \right) - \mu_i^{(q)} \mu_i^{(q)*} \right) \\
&= \left(\langle [\boldsymbol{a}_k]_q^{j-i} \rangle \right)^* \text{var}([\boldsymbol{\alpha}_k]_q),
\end{aligned} \tag{25}$$

where $\mathbb{E}\left([\boldsymbol{\psi}_k^{(q)}]_i \tilde{\psi}'^* \right) - \mu_i^{(q)} \tilde{\mu}^* = 0$ because of the definition of $\tilde{\psi}$ as independent of $[\boldsymbol{\psi}_k^{(q)}]_i$. An equivalent derivation can be made for terms in the lower half of the covariance matrix, thus we have derived the covariance matrix shown in (21). As the total received signal corresponds to the sum of the received signal for all the targets, this translates into a circularly symmetric complex Gaussian whose covariance matrix is the sum of the covariance matrices of all the targets, as indicated in (20), which thus yields (19).

The key property of this construction of the covariance matrix is that the MIMO physics is implicitly encoded in the point estimate of $\langle [\boldsymbol{a}_k]_q^{j-i} \rangle$, which distinguishes our method from general linear-Gaussianisation of a non-linear-Gaussian problem. Furthermore, we can physically reason about the

$$\frac{\partial \Sigma(\mathbf{S}'_k)}{\partial (\mathbf{S}'_k)_{i,j}} = \mathrm{tr}\left\{\sigma_n^{-2}(\sigma_n^{-2}\mathbf{S}_k'^H\mathbf{S}'_k + \mathbf{R}_k^{-1})^{-1}((\mathbf{J}^{(i,j)})^H\mathbf{S}'_k + \mathbf{S}_k'^H\mathbf{J}^{(i,j)})(\sigma_n^{-2}\mathbf{S}_k'^H\mathbf{S}'_k + \mathbf{R}_k^{-1})^{-1}\right\}, \tag{29}$$

form of the covariance matrix. For we can see that $\langle [\boldsymbol{a}_k]_q^i \rangle \triangleq \mathbb{E}([\boldsymbol{a}_k]_q^i | \mathbf{X}^{k-1}) \neq \exp(2\pi\sqrt{-1}(\tilde{\lambda}/\lambda)\sin(\mathbb{E}([\boldsymbol{\phi}_k]_q^i)))$. This property enables us to make a qualitative justification for this method. Given that the $[\boldsymbol{a}_k]_q^i$ has support only on the unit circle, $|\langle [\boldsymbol{a}_k]_q^i \rangle| \leq 1$ (for all $\langle a_q \rangle$), which is a required property for the covariance matrix to be valid. Moreover, we can interpret this property that, as we gain more knowledge about the angle ϕ, then $|\langle [\boldsymbol{a}_k]_q^i \rangle|$ becomes closer to one and the correlation between the elements of $\boldsymbol{\psi}_k^{(q)}$ becomes stronger, as we would expect from physical reasoning. Finally, we can see that, as i increases, $|\langle [\boldsymbol{a}_k]_q^i \rangle|$ decreases, which means that more separated powers of $\langle [\boldsymbol{a}_k]_q^i \rangle$ are treated as less correlated. That is, more separated elements are treated as less correlated, which is consistent with what we would expect from physical reasoning including spatial diversity.

It is also worth noting that, throughout this section we have used simplified notation for the complex Gaussian, as the relation matrix is always zero (the circularly symmetric scenario), for our method this property is upheld throughout the successive reconstruction of $\mathbf{R}_k$, regardless of $p(\boldsymbol{\theta}_k|\mathbf{X}^{k-1})$.

D. Relating the cost function of $\boldsymbol{\psi}$ to that of $\boldsymbol{\theta}$

The purpose of the above re-formulation is that the MSE of $\boldsymbol{\psi}_k$ can be minimised, rather than that of $\boldsymbol{\theta}_k$ directly. Intuitively, given that $\boldsymbol{\theta}_k$ is a deterministic function of $\boldsymbol{\psi}_k$, it seems reasonable that designing the beam-pattern to reduce the trace of the expected covariance matrix of the estimate of $\boldsymbol{\psi}_k$ should reduce the trace of the expected covariance matrix of the estimate of $\boldsymbol{\theta}_k$. More formally, the trace of the variance of $\boldsymbol{\theta}_k$ can be expressed (by definition):

$$\mathrm{tr}\left\{\mathrm{covar}\left(\boldsymbol{\theta}_k\right)\right\} = \sum_{q=1}^{Q'} \mathrm{var}([\boldsymbol{\alpha}_k]_q) + \mathrm{var}([\boldsymbol{\phi}_k]_q), \tag{26}$$

whereas from (21):

$$\mathrm{tr}\left\{\mathbf{R}_k\right\} = (N_T + N_R - 1)\sum_{q=1}^{Q'} \mathrm{var}([\boldsymbol{\alpha}_k]_q). \tag{27}$$

From (26) and (27), we can see that minimising the trace of the covariance matrix associated with the estimate of $\boldsymbol{\psi}_k$ only minimises the variance of the target attenuation, and not the target angle. However, the non-diagonal terms have been artificially constructed to account for the physics of the MIMO array and encode information about the target angles (as described in Section III-C), and these *are* a contributing factor in the cost function formulation (to be given in Section IV), and therefore optimisation of the cost function implicitly estimates the target angles. In simple terms, the non-diagonal terms dictate the correlations between the target attenuations, which depend on the respective target angles. Therefore designing the waveform such that uncertainty of the target angles (and therefore the correlations between the target attenuations) is reduced will in turn reduce the uncertainty of the target attenuations, as is the explicit aim of the cost function optimisation.

In Appendix A we show that $\mathbf{R}_k$ can be constructed for the case where the element spacing is not the same for the transmit and receive arrays (and indeed, need not be equal within each array), therefore the following optimisation applies to the general MIMO active sensing system case.

IV. Optimisation of linear-Gaussian approximate cost function

Having expressed the state-space model as a linear function of $\boldsymbol{\psi}_k$, as given in (10) (i.e., as opposed to the prior, non-linear function of $\boldsymbol{\theta}_k$), and argued that designing a beam-pattern to minimise the mean squared error of $\boldsymbol{\psi}_k$ will have the effect of minimising the mean squared error of $\boldsymbol{\theta}_k$, it remains to formulate this minimisation problem. This can be achieved by applying the analysis leading to [9, Eq. (13)], to our problem. Using our nomenclature we formulate the proposed approximate optimisation as the minimisation of the cost function, $\Sigma(\mathbf{S}'_k)$, subject to maximum power P:

$$\begin{aligned} &\min_{\mathbf{S}'_k} \ \Sigma(\mathbf{S}'_k) = \mathrm{tr}\left\{(\sigma_n^{-2}\mathbf{S}_k'^H\mathbf{S}'_k + \mathbf{R}_k^{-1})^{-1}\right\} \\ &\text{s.t.} \quad \frac{1}{L}\mathrm{tr}\left\{\mathbf{S}_k^H\mathbf{S}_k\right\} \leq \mathrm{P} \end{aligned} \tag{28}$$

Owing to the constraints on the construction of $\mathbf{S}'_k$, described in Section III-B, we cannot use the additional analysis detailed by Yang and Blum [9] to find $\mathbf{S}'_k$ directly. It is also the case that this is not a convex optimisation [23]. However, it is possible to express the gradient of Σ with respect to the (i,j)th element of $\mathbf{S}'_k$ (for any element not constrained to be equal to zero in the construction of $\mathbf{S}'_k$, as defined in Section III-B). This is shown in (29), in which $\mathbf{J}^{(i,j)}$ is a matrix of zeros except for a single entry of 1 at (i,j). Let the elements of $\mathbf{S}'_k$ not constrained to be zero be stacked in a vector $\mathbf{s}'_k$, such that

$$\mathbf{s}'_k = \mathbf{B}\mathrm{vec}(\mathbf{S}_k), \tag{30}$$

where $\mathbf{B} \in \mathbb{R}^{LN_TN_R \times LN_T}$ ($\mathbf{B}$ is fully defined in Appendix B), thus the dimensions of $\mathbf{B}$ are such that we can express:

$$\mathrm{vec}(\mathbf{S}_k) = \mathbf{B}^\dagger\mathbf{s}'_k, \tag{31}$$

where $\mathbf{B}^\dagger = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$, i.e., $(.)^\dagger$ denotes the pseudo-inverse. So it follows that:

$$\nabla_{\mathrm{vec}(\mathbf{S}_k)}(\Sigma) = \mathbf{B}^\dagger\nabla_{\mathbf{s}'_k}(\Sigma), \tag{32}$$

where the elements of $\nabla_{\mathbf{s}'_k}(\Sigma)$ can be calculated according to (29), thus enabling us to express the gradient of Σ with respect to the elements of $\mathbf{S}_k$. In Appendix B we show that the structure of $\mathbf{B}$ is such that the elements of $\nabla_{\mathrm{vec}(\mathbf{S}_k)}(\Sigma)$ can be calculated using fewer operations than performing the matrix pseudo-inverse – both for when the transmit and receive

Algorithm 1 Simple pseudo-code for optimisation of the linear-Gaussian approximation

Initialise: $p(\boldsymbol{\theta}_0)$
For: $k = 1 : K$
from $p(\boldsymbol{\theta}_k|\mathbf{X}^{k-1})$ determine $\mathbf{R}_k$ according to (20)
use gradient descent to design $\mathbf{S}_k$ (28)
transmit $\mathbf{S}_k$
receive $\mathbf{X}_k$
determine $p(\boldsymbol{\theta}_{k+1}|\mathbf{X}_k)$

arrays have the same spacing, and for when they do not.

This enables (28) to be optimised using gradient descent, as detailed in Algorithm 1. As in [5], it is possible to handle the power constraint by descending in the direction of the component of the gradient perpendicular to $\text{vec}(\mathbf{S}_k)$ and subsequently renormalising (assuming that the power constraint is always satisfied with equality):

$$\nabla^{\perp}_{\text{vec}(\mathbf{S}_k)}(\Sigma) = \nabla_{\text{vec}(\mathbf{S}_k)}(\Sigma) - \text{vec}(\mathbf{S}_k)\frac{(\nabla_{\text{vec}(\mathbf{S}_k)}(\Sigma))^T\text{vec}(\mathbf{S}_k)}{(\text{vec}(\mathbf{S}_k))^T\text{vec}(\mathbf{S}_k)}. \quad (33)$$

V. Main results

We present some results to demonstrate that the proposed ABD method does indeed lead to a reduction in root mean squared error (RMSE) compared to the non-adaptive case, in which the transmit beam-pattern is such that the same power is transmitted at all angles. To do so, we simulate a MIMO active sensing system with half-wavelength spacing on both the transmit and receive arrays and $N_T = N_R = 7$, we also set $L = 7$. The MIMO active sensing system must estimate both the target attenuation and the target angle. We simulate three scenarios: a single moving target; two moving targets; and two static targets, and in each case consider the first 20 steps.

A. Single moving target

The single target was initially located at $\phi_1 = -50^o$, and thereafter moved in a random walk: that is, $\phi_{k+1} \sim \mathcal{N}(\phi_k, \sigma^2_\phi)$, for which set the standard deviation $\sigma_\phi = 0.5^o$, and $\arg(\boldsymbol{\alpha})$ and $|\boldsymbol{\alpha}|$ did not vary. The statistical definition of the random walk was available to the MIMO active sensing system (i.e., there was no model mismatch). We set the array signal-to-noise ratio (ASNR) at 3 dB, where ASNR $\triangleq |\alpha|^2 \text{P} N_R L/(0.5\sigma^2_n)$ (in which the factor 0.5 in the denominator is introduced owing to our definition of σ^2_n as the noise variance for each of the real and imaginary components).

The underlying estimation of $\boldsymbol{\theta}_k$ was conducted by a PF with 6120 particles, initially placed on a grid with resolution 10^o for both ϕ and $\arg(\alpha)$ and 0.3 for $|\alpha|$ out to 3 (the noise was set such that $\alpha = 1$). The particles were resampled at each step of the PF, and then each new particle was randomly perturbed such each of $\arg(\alpha)$, $|\alpha|$ and ϕ were moved to a new location according to a normal random process whose mean was the previous location, and whose standard deviation was $5 \times 0.85^{k-1}$ for both $\arg(\alpha)$ and ϕ, and $0.15 \times 0.85^{k-1}$ for $|\alpha|$. The rationale behind this operation of the PF was to allow a relatively small number of particles to ultimately cover the entire support of the PDF. The factor $5 \times 0.85^{k-1}$ was used to reduce the variance throughout the simulated trials, with the aim of ultimately attaining an accurate estimation of the target parameters.

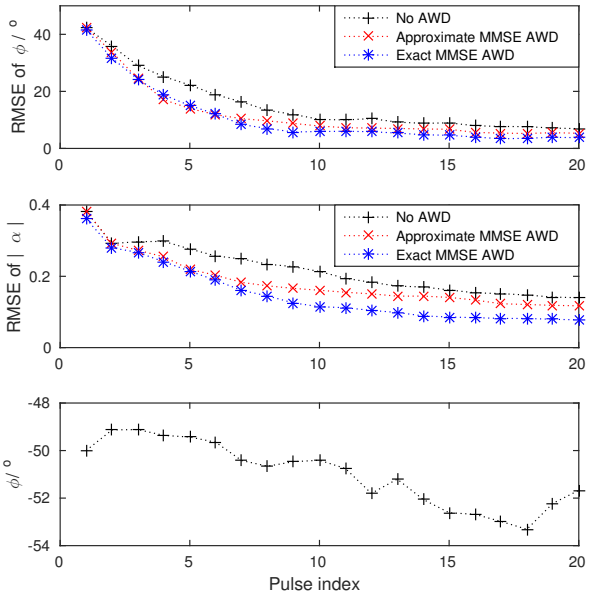

Fig. 1. RMSE for a single moving target (averaged over 500 trials) for no ABD, approximate MMSE ABD and exact MMSE ABD.

For the single moving target simulation, we also include results for the exact MMSE ABD method [5] with the same PF and gradient descent set-up as for the approximate MMSE ABD method. In addition to the PF particles, the exact MMSE ABD method also samples from the PF a number of times, N_S, which we set to be 250. Unlike for the approximate MMSE ABD method proposed herein, for the exact MMSE ABD it was necessary to weight the relative importance of the error in the estimation of the target angle as well as the argument and magnitude of the target attenuation. This was done such that the target angle and target attenuation magnitude had approximately equal weighting and zero weighting was given to the argument of the target attenuation. This set-up is physically justified as the target attenuation magnitude may be useful in practise, for example indicating range, whereas the argument of the target attenuation is unlikely to be of interest. We also include results for the non-adaptive case, again with the same PF set-up. The results are shown in Fig. 1, with the RMSE approximated by averaging over 500 trials. In Section V-D we discuss the results.

B. Two moving targets

We also include results for two target parameter estimation (estimation of $\boldsymbol{\psi}$, $\arg(\boldsymbol{\alpha})$ and $|\boldsymbol{\alpha}|$), for a scenario, ASNR=3 dB where $|\boldsymbol{\alpha}_1| = |\boldsymbol{\alpha}_2|$ (i.e., the two targets reflect the same

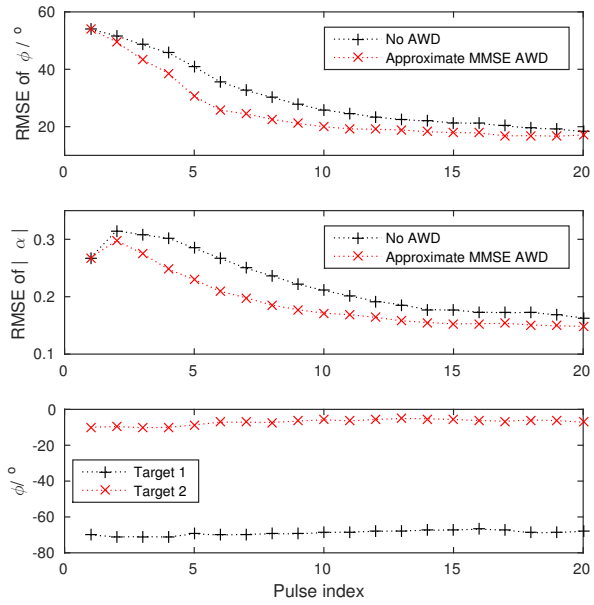

Fig. 2. RMSE for two moving targets (averaged over 500 trials) for no ABD, approximate MMSE ABD and exact MMSE ABD.

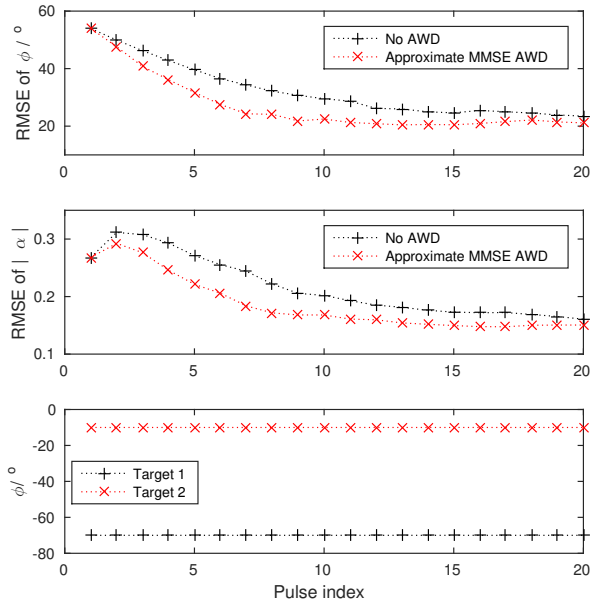

Fig. 3. RMSE for two static targets (averaged over 500 trials) for no ABD, approximate MMSE ABD and exact MMSE ABD.

magnitude of power).

Such was the high dimensionality of the two target problem, that is was no longer feasible to initialise the particles on a grid, and instead $N_P = 10000$ particles, were placed at random. Additionally, we implemented a slightly different resampling policy: when the number of effective particles ($N_{eff} = 1/\sum_{i=1}^{N_P}(w^{(i)})^2$, see [24, ch. 3]) fell below $0.2 \times N_P$. Once again to increase the diversity of the particle locations, we took the resampled particle parameter values not as the values of the new particles themselves, but as the mean of a random variable, drawn from a normal distribution, where the standard deviation was 5^o for ϕ and $\arg(\alpha)$ and 0.01 for $|\alpha|$ (again the ASNR was set such that $|\boldsymbol{\alpha}| = 1$). Unlike for the one target scenario, we found that it was not beneficial to reduce these standard deviations throughout using the factor 0.85^{k-1}.

We simulated a scenario with moving targets, with initial location $\boldsymbol{\phi}_1 = [-70; -10]$, evolving as a random walk with each target having the same standard deviation as in the single target case, i.e., $\boldsymbol{\phi}_{k+1} \sim \mathcal{N}\left(\boldsymbol{\phi}_k, \sigma_\phi^2 \mathbf{I}_2\right)$, and $\arg(\boldsymbol{\alpha})$ and $|\boldsymbol{\alpha}|$ did not vary. Fig. 2 shows the RMSE for this set-up, again found by averaging over 500 trials.

C. Two static targets

For the two static target set-up we used the same simulation set-up as that for two moving targets, except that we set $\boldsymbol{\phi}_k = [-70; -10]$ for $1 \leq k \leq 20$. The results for the RMSE (averaged over 500 trials) are shown in Fig. 3. We also show an example of a beam-pattern designed by the approximate MMSE ABD method we propose, shown in Fig. 4 (in which the transmit power, shown on the vertical axis, is defined as $(1/L)\mathbf{a}_T^H(\phi)(\mathbf{S}_k^*\mathbf{S}_k^T)\mathbf{a}_T(\phi)$).

D. Discussion of results

The results show that the proposed approximate MMSE ABD method improves the RMSE performance in comparison to the no ABD case, but not as much as the exact MMSE ABD method, as would be expected. Note that we do not compare to the method proposed by Huleihel *et al* [4], as this had inferior performance compared to the exact MMSE ABD method for similar computational complexity [5]. For the multiple target tracking, such was the high dimensionality of the estimation that it was not possible to set-up the PF to yield an good ultimate convergence of the target parameters. Thus, in the results presented, we have used a relatively low ASNR to highlight the improved performance in the early stages of convergence. It should, however, be noted that there exists a large body of existing work on how to adapt PFs to handle adverse estimation scenarios (for example [25]–[27]), but detailed examination thereof is outside of the scope of this paper. It is important to note that, as long as the conditions to avoid under-determination set-out in Section III-A are adhered to and the number of particles in the PF is sufficiently large, then there is no reason in principle why the proposed method should not yield good target parameter estimation for scenarios where there is a large number of targets.

VI. Computational complexity

It is possible to quantify the reduction in computational complexity of this proposed method by considering how the number of floating point operations required varies as a function of the generalised set of parameters that defines the operational set-up.

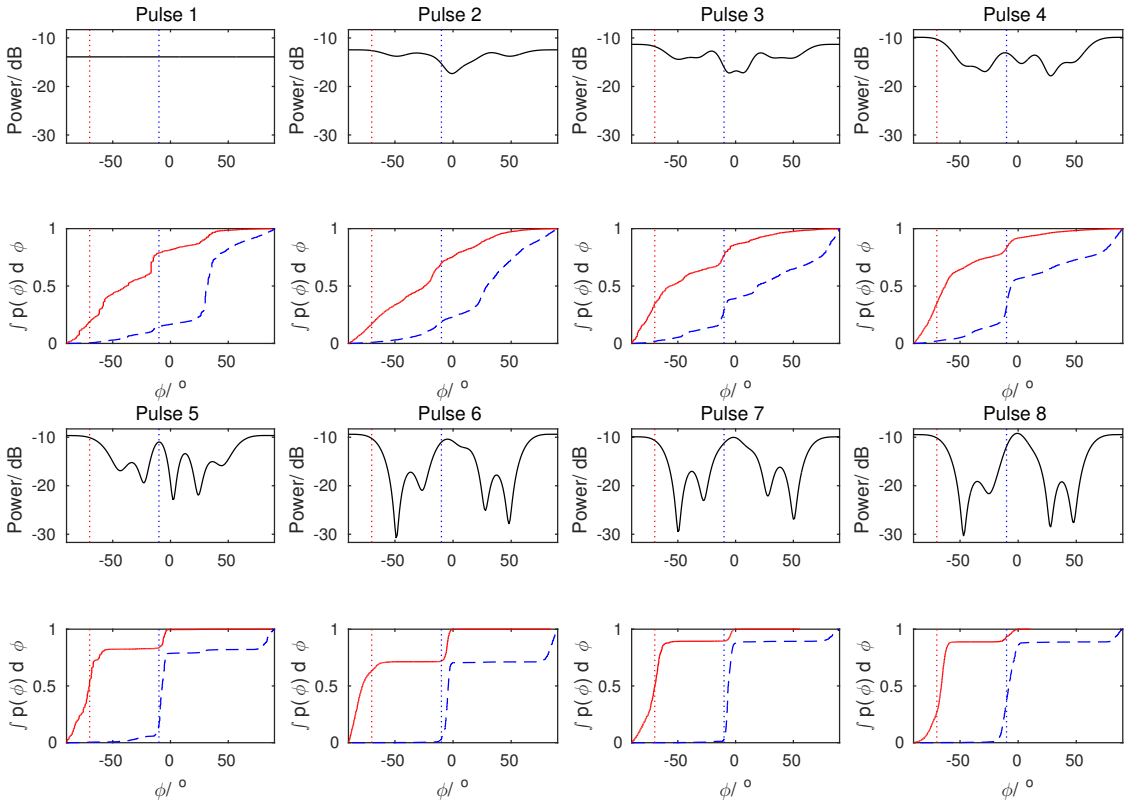

Fig. 4. Example of approximate MMSE ABD for the first 8 pulses of the scenario with two static targets. The second and fourth rows correspond to cumulative density functions, and the vertical dotted lines indicate the actual target locations.

Task	**Number of operations**	
	Method in [5]	Method proposed herein
Draw N_S samples	$\mathcal{O}(N_S)$	–
Construct $\mathbf{R}_k$	–	$\mathcal{O}(QN_p(N_T+N_R))$
Evaluate cost function	$\mathcal{O}(N_c N_S^2(Q+LN_TN_R))$	$\mathcal{O}(N_c((N_T+N_R)^3+LN_R(N_T+N_R)^2))$
Evaluate derivative of cost function	$\mathcal{O}(N_d N_S^2(LN_TQ+L^2N_T^2N_R))$	$\mathcal{O}(N_d((N_T+N_R)^3+LN_R(N_T+N_R)^2))$

TABLE I
COMPUTATIONAL COMPLEXITY AS A FUNCTION OF ALL PARAMETERS

A. Complexity as a function of all parameters

There are three functions that contribute to the overall computational load, which are required at each step: construction of $\mathbf{R}_k$; evaluation of the cost function (by definition, N_c times per step); and evaluation of the derivative of the cost function (by definition, N_d times per step).

1) Construction of $\mathbf{R}_k$ requires $\mathcal{O}(Q)$ sums over N_p particles for each of $\mathcal{O}(N_T+N_R)$ powers of $\langle[\boldsymbol{a}_k]_q\rangle$, resulting in an overall computational complexity $\mathcal{O}(QN_p(N_T+N_R))$.
2) Evaluation of the cost function computational load is dominated by the inversion of $\mathbf{R}_k$, which requires $\mathcal{O}((N_T+N_R)^3)$ operations and the matrix multiplication $\mathbf{S}_k'^H\mathbf{S}_k'$, which requires $\mathcal{O}(LN_R(N_T+N_R)^2)$ operations, leading to an overall computation complexity $\mathcal{O}((N_T+N_R)^3+LN_R(N_T+N_R)^2)$.
3) The computational load of the evaluation of the derivative of the cost function is dominated by the same terms as those for the cost function.

This information, along with the corresponding terms in [5] is presented in Table I, which clearly illustrates the primary computational saving of the proposed method: that, unlike the exact MMSE ABD method [5], evaluation of the cost function and its derivative (each potentially required a large number of times per step) is *not* a function of the number of samples/particles (which is likely to be large). It should, however, be noted that the more sophisticated algebra of the method proposed herein manifests itself as a greater complexity associated with the number of array elements. For simplicity of exposition let $N_T = N_R = N_{TR}$, then we the cost function evaluation grows as $\mathcal{O}(N_{TR}^3)$ rather than $\mathcal{O}(N_{TR}^2)$ (the derivative evaluation grows as $\mathcal{O}(N_{TR}^3)$ for both methods). Whilst, to reiterate, this is not a significant factor for the types of problems we consider, it may become so if

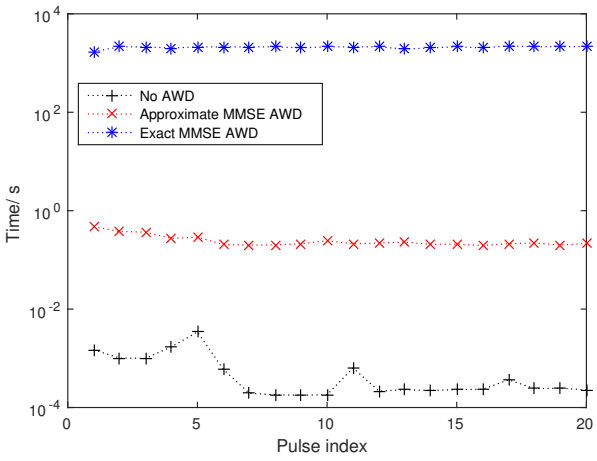

Fig. 5. Computation time for ABD in single target scenario.

Task	Number of operations	
	Method in [5]	Method proposed herein
Draw N_S samples	$\mathcal{O}(\exp(Q))$	–
Construct $\mathbf{R}_k$	–	$\mathcal{O}(\exp(Q))$
Evaluate cost function	$\mathcal{O}(N_c \exp(Q))$	$\mathcal{O}(N_c Q^3)$
Evaluate of derivative	$\mathcal{O}(N_d \exp(Q))$	$\mathcal{O}(N_d Q^3)$

TABLE II
COMPUTATIONAL COMPLEXITY AS A FUNCTION OF THE NUMBER OF TARGETS

this method is applied to very large MIMO arrays, thus it is worth explicitly raising this point.

To further demonstrate the computational complexity reduction of the approximate method we propose here, we include an example of the computation time for the approximate method and the exact MMSE ABD method [5]. For this we performed a single run of each method using *Matlab* on an Intel Core 1.9 GHz i3 4030U processor with memory of 4 GB 1600 MHz DDR3L SSDRAM. For comparison we also show the computation time for the non-adaptive case, where just the computation required for updating the PF is required. Fig. 5 shows the computation times for the three methods. The approximate MMSE ABD method clearly consumes much less time than the exact MMSE ABD method: the former has an average computational time per step of 1.99×10^3 s, whereas the latter has an average computational time per step of 0.237 s, i.e., a factor of 8.40×10^4 reduction in computation time. Such a large computation time saving is not unexpected, as the primary purpose of the approximate MMSE ABD method proposed herein is to remove the need for computationally expensive summing over particles.

B. Complexity as a function of number of targets

Rather than independently setting all of the MIMO active sensing system parameters, in reality the system will be designed to a specification that includes a maximum number of targets to be tracked. Therefore it is of interest to express the computational complexity of the method proposed herein, compared to the exact method [5], as a function of the number of targets parameters, Q (which is proportional to the number of targets, Q'). To do this, we need to express the MIMO active sensing system parameters in terms of their relationship with the number of target parameters. As noted in Section III-A, $N_T + N_R$ must grow linearly with Q' so that the system of linear equations is not under-determined. In Section III-B we note that L need not grow with the number of target parameters. For a specified resolution, the number of particles, N_P, in the PF must grow exponentially with the number of targets (i.e., each additional target parameter to be estimated introduces a new dimension into the PF estimation, which therefore requires a multiplicative factor increase in the number of targets to retain the same resolution for estimation of the target parameters). In order to retain ABD performance, the number of samples, N_S must grow at least as fast as the number of particles. The resultant computational complexity expressions are shown in Table II.

Table II shows that for both methods there is an unavoidable complexity growth with the number of targets to prepare the cost function, drawing samples for the exact method [5] and constructing $\mathbf{R}_k$ for the method proposed herein. This is not necessarily a problem, as this only occurs once, and given its nature as preparation for the actual cost function optimisation, the cost can be absorbed into the PF itself (whose complexity grows with the number of particles, by definition). Of more concern is the complexity associated with the cost function and derivative evaluations, which will typically be conducted many times in the optimisation process. For these, we can see that the exact method has a computational complexity that grows exponentially with the number of targets, whereas the for approximate method the complexity is polynomial. This represents a distinct advantage of the approximate method for scenarios where a large number of target parameters are to be estimated.

One consequence of this reduction of complexity class from exponential to polynomial is that there exists scenarios where the exact method has both longer run-time and worse MMSE performance than the approximate method. This is because, for large N_c and N_d the computational complexity associated with the exact method will always exceed that of the approximate method for sufficiently large Q, even if we reduce the required resolution (i.e., average concentration of particles) to such an extent that the PF doesn't function properly, and therefore attempting ABD is futile. To give an example of this, which is somewhat imprecise but nevertheless suffices to illustrate the principle, consider that we could specify that the required PF particle resolution corresponds to particles initially placed on a Q dimensional grid, with two particles per target parameter, therefore $N_P = 2^Q$. In such a case, the number of particles would be insufficient for the PF (and therefore ABD) to function effectively (apart from perhaps after many iterations, if the PF particular re-distribution method is very good), even though the computational complexity will eventually exceed that of the approximate method, with an appropriate number of particles (and therefore with the ABD working properly), as Q is increased (i.e., because exponential will always exceed polynomial).

VII. Conclusions

In this paper, we have expressed the cost function for MMSE ABD in MIMO active sensing systems as a linear-Gaussian approximation that enables computationally simple optimisation. Analysis of the computational complexity indicates the potential of the proposed method to reduce the computation time of MMSE ABD, most significantly that the number of operations required grows only polynomially, rather than exponentially (as is the case for the exact method) with the number of targets being tracked. Numerical results support this theoretical reduction in computational complexity, showing that the proposed approximate MMSE ABD method leads to a reduction in RMSE compared to the no ABD case, and a four orders of magnitude reduction in computation time compared to the exact MMSE ABD method (which slightly out-performs the proposed approximate MMSE ABD method in terms of RMSE performance). The proposed approximate MMSE ABD method relies on both the target angle and attenuation being estimated, which is a more general case compared to that considered in the numerical results in the paper in which the exact MMSE ABD method was proposed [5]. Therefore the numerical results included herein, for both the exact and approximate MMSE ABD methods estimate the target attenuation as well as angle, demonstrating good performance of each in this more general, and physically reasonable set-up.

The results we present in this paper are for the case where the elements in transmit and receive arrays have the same uniform spacing, however we present analysis demonstrating that the proposed approximate ABD method applies to MIMO adaptive sensing systems in general, regardless of array spacing. Therefore an important extension to this work is to apply the method to MIMO adaptive sensing systems with unequal/ non-uniform arrays.

Another extension concerns the fact that the analysis in this paper implicitly treats all reflective objects as targets, however it may be the case that motion analysis reveals some of these reflective objects to be clutter rather than targets. In such a scenario, it would be interesting to introduce additional constraints in the MMSE optimisation formulation such that the transmitted power in the direction of the clutter does not exceed some defined threshold, thus achieving joint clutter suppression and MMSE ABD.

Appendix A
Construction of $\mathbf{R}_k$ for unequal and/ or non-uniform array spacing

The analysis in Section III, leading to the approximate ABD method relies on two features of ABD in MIMO active sensing systems: firstly, that the elements of $\mathbf{H}^{(q)}$ can also be expressed as the same complex number to a power as in (6); and secondly that the system model can be re-written such that the unique elements of $\mathbf{H}$ are expressed as a vector, that is then pre-multiplied by a matrix where each element is equal to zero or an element of $\mathbf{S}_k$, as in (9). From the definitions of steering vectors, it is trivial to see that the first of these features is common to all such problems – if the transmit and receive arrays are unequal and/or non-uniform fractional powers may be required, but that would not invalidate the ABD method.

For the second feature, when $\mathbf{H}^{(q)}$ is such that *all* of the elements therein are distinct, then $\text{vec}(\mathbf{H})$ is equivalent to $\boldsymbol{\psi}_k$ (i.e., each is a vector of the distinct elements of $\mathbf{H}$ for its respective system model). In this case, dropping the subscript k, we have that:

$$\begin{aligned}\mathbf{X} &= \mathbf{HS} + \mathbf{N} \\ &= \mathbf{I}_{N_R}\mathbf{HS} + \mathbf{N},\end{aligned} \tag{34}$$

where $\mathbf{I}_{N_R}$ is the identity matrix of size N_R. From (34) we can express the vectorised form:

$$\begin{aligned}\text{vec}(\mathbf{X}) &= \text{vec}(\mathbf{I}_{N_R}\mathbf{HS}) + \text{vec}(\mathbf{N}) \\ &= \mathbf{S}^T \otimes \mathbf{I}_{N_R}\text{vec}(\mathbf{H}) + \text{vec}(\mathbf{N}),\end{aligned} \tag{35}$$

from [28], which is the desired form with a simple function of $\mathbf{S}$ where each element is equal to either zero or an element of $\mathbf{S}$ (in this case $\mathbf{S}^T \otimes \mathbf{I}_{N_R}$) pre-multiplying a vectorised version of $\mathbf{H}$.

Appendix B
Pseudo-inversion of $\mathbf{B}$

Noticing that each non-zero element of $\mathbf{S}'_k$ is equal to exactly one of the elements of $\mathbf{S}_k$, according to the definition in (10), without loss of generality we can re-order $\text{vec}(\mathbf{S}_k)$ and $\mathbf{s}'_k$ (denoted $\tilde{\mathbf{s}}_k$ and $\tilde{\mathbf{s}}'_k$ respectively) such that:

$$\tilde{\mathbf{s}}'_k = \tilde{\mathbf{B}}\tilde{\mathbf{s}}_k, \tag{36}$$

where $\tilde{\mathbf{B}} = [\mathbf{I}^{(1)}_{LN_T}, \ldots, \mathbf{I}^{(N_R)}_{LN_T}]^T$, i.e., N_R identities vertically stacked. This enables us to express:

$$\begin{aligned}\tilde{\mathbf{B}}^\dagger &= (\tilde{\mathbf{B}}^T\tilde{\mathbf{B}})^{-1}\tilde{\mathbf{B}}^T \\ &= \left([\mathbf{I}^{(1)}_{LN_T}, \ldots, \mathbf{I}^{(N_R)}_{LN_T}] \begin{bmatrix} \mathbf{I}^{(1)}_{LN_T} \\ \vdots \\ \mathbf{I}^{(N_R)}_{LN_T} \end{bmatrix} \right)^{-1} [\mathbf{I}^{(1)}_{LN_T}, \ldots, \mathbf{I}^{(N_R)}_{LN_T}] \\ &= (N_R\mathbf{I}_{LN_T})^{-1}[\mathbf{I}^{(1)}_{LN_T}, \ldots, \mathbf{I}^{(N_R)}_{LN_T}] \\ &= \frac{1}{N_R}[\mathbf{I}^{(1)}_{LN_T}, \ldots, \mathbf{I}^{(N_R)}_{LN_T}],\end{aligned} \tag{37}$$

which allows us to calculate the elements of $\nabla_{\text{vec}(\mathbf{S}_k)}(\Sigma)$ in (32) by taking the average of corresponding elements of $\nabla_{\text{vec}(\mathbf{S}'_k)}(\Sigma)$. Consider the example in (12), we can see that $s_{1,1}$ appears three times in $\mathbf{S}'_k$ (at s'_{11}, s'_{22} and s'_{33}), so we would calculate the differential of the cost with respect to $s_{1,1}$ by summing one third of the differential of the cost with respect to s'_{11}, s'_{22} and s'_{33}, as found in (29).

References

[1] S. Haykin, "Cognitive radar: a way of the future," *IEEE Signal Processing Magazine*, vol. 23, no. 1, pp. 30–40, Jan 2006.

[2] J. Ender and S. Brüggenwirth, "Cognitive radar - enabling techniques for next generation radar systems," in *2015 16th International Radar Symposium (IRS)*, June 2015, pp. 3–12.

[3] S. P. Sira, Y. Li, A. Papandreou-Suppappola, D. Morrell, D. Cochran, and M. Rangaswamy, "Waveform-agile sensing for tracking," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 53–64, Jan 2009.

[4] W. Huleihel, J. Tabrikian, and R. Shavit, "Optimal adaptive waveform design for cognitive MIMO radar," *IEEE Transactions on Signal Processing*, vol. 61, no. 20, pp. 5075–5089, Oct 2013.

[5] S. Herbert, J. Hopgood, and B. Mulgrew, "MMSE adaptive waveform design for active sensing with applications to MIMO radar," *IEEE Transactions on Signal Processing*, vol. PP, no. 99, pp. 1–1, 2017.

[6] N. Sharaga, J. Tabrikian, and H. Messer, "Optimal cognitive beamforming for target tracking in MIMO radar/sonar," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 8, pp. 1440–1450, Dec 2015.

[7] S. Herbert, J. R. Hopgood, and B. Mulgrew, "Optimality criteria for adaptive waveform design in MIMO radar systems," in *2017 Sensor Signal Processing for Defence Conference (SSPD)*, Dec 2017, pp. 1–5.

[8] A. Donev, A. Atkinson, and R. Tobias, *Optimum Experimental Designs, with SAS*, ser. Oxford Statistical Science Series. United Kingdom: Oxford University Press, 2007.

[9] Y. Yang and R. S. Blum, "MIMO radar waveform design based on mutual information and minimum mean-square error estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 1, pp. 330–343, January 2007.

[10] H. Sorenson, *Kalman Filtering: Theory and Application*, ser. IEEE Press selected reprint series. IEEE Press, 1985. [Online]. Available: https://books.google.co.uk/books?id=2pgeAQAAIAAJ

[11] S. J. Julier and J. K. Uhlmann, "A new extension of the kalman filter to nonlinear systems," vol. 3068, 02 1999.

[12] E. A. Wan and R. V. D. Merwe, "The unscented kalman filter for nonlinear estimation," in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*, 2000, pp. 153–158.

[13] S. Herbert, J. Hopgood, and B. Mulgrew, "MMSE adaptive waveform design for a MIMO active sensing system tracking multiple moving targets," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2018.

[14] F. Gini, A. D. Maio, and L. Patton, Eds., *Waveform Design and Diversity for Advanced Radar Systems*, ser. Radar, Sonar &; Navigation. Institution of Engineering and Technology, 2012. [Online]. Available: http://digital-library.theiet.org/content/books/ra/pbra022e

[15] J. Liu, H. Li, and B. Himed, "Joint optimization of transmit and receive beamforming in active arrays," *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 39–42, Jan 2014.

[16] P. Stoica, H. He, and J. Li, "Optimization of the receive filter and transmit sequence for active sensing," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1730–1740, April 2012.

[17] A. Aubry, A. D. Maio, and M. M. Naghsh, "Optimizing radar waveform and doppler filter bank via generalized fractional programming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 8, pp. 1387–1399, Dec 2015.

[18] F. Hlawatsch, "Range-doppler estimation," in *Time-Frequency Analysis and Synthesis of Linear Signal Spaces*, ser. The Kluwer International Series in Engineering and Computer Science. Springer, 1998, vol. 440.

[19] I. Bekkerman and J. Tabrikian, "Target detection and localization using MIMO radars and sonars," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3873–3883, Oct 2006.

[20] T. Riedrich, "Partington, j. r., an introduction to hankel operators. cambridge etc., cambridge university press 1988. 103 pp., 20.00 h/b, 7.50 p/b. isbn 0-521-36791-3 (london mathematical society student texts 13)," *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift fr Angewandte Mathematik und Mechanik*, vol. 71, no. 11, pp. 464–464, 1991. [Online]. Available: http://dx.doi.org/10.1002/zamm.19910711112

[21] R. Gallager, *Stochastic Processes: Theory for Applications*, ser. Stochastic Processes: Theory for Applications. Cambridge University Press, 2013. [Online]. Available: https://books.google.co.uk/books?id=ERLrAQAAQBAJ

[22] E. L. Melnick and A. Tenenbein, "Misspecifications of the normal distribution," *The American Statistician*, vol. 36, no. 4, pp. 372–373, 1982. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/00031305.1982.10483052

[23] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[24] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, 2003. [Online]. Available: https://books.google.co.uk/books?id=zABIY–qk2AC

[25] T. Li, M. Bolic, and P. M. Djuric, "Resampling methods for particle filtering: Classification, implementation, and strategies," *IEEE Signal Processing Magazine*, vol. 32, no. 3, pp. 70–86, May 2015.

[26] I. Kyriakides, T. Trueblood, D. Morrell, and A. Papandreou-Suppappola, "Multiple target tracking using particle filtering and adaptive waveform design," in *2008 42nd Asilomar Conference on Signals, Systems and Computers*, Oct 2008, pp. 1188–1192.

[27] L. Miao, J. J. Zhang, C. Chakrabarti, and A. Papandreou-Suppappola, "A new parallel implementation for particle filters and its application to adaptive waveform design," in *2010 IEEE Workshop On Signal Processing Systems*, Oct 2010, pp. 19–24.

[28] K. M. Abadir and J. R. Magnus, *Matrix Algebra*, ser. Econometric Exercises. Cambridge University Press, 2005.

Steven Herbert received the M.A. (cantab) and M.Eng. degrees from the University of Cambridge Department of Engineering in 2010, and the Ph.D. degree from the University of Cambridge Computer Laboratory in 2015. Following the submission of his Ph.D. thesis he continued to work at the University of Cambridge Computer Laboratory as a Research Assistant from May to June 2014, and remained a visiting researcher until December 2016. From January 2015 until July 2016 he worked at Blu Wireless Technology, Bristol, during which time he was a co-inventor of five patents. From September 2016 until January 2018 he was a Research Associate at the University of Edinburgh School of Engineering, where he worked on adaptive beam-pattern design for active sensing, and he is currently a visiting researcher at the University of Edinburgh. He started his current position, Research Associate at the University of Cambridge Department of Applied Mathematics and Theoretical Physics in February 2018, where he is working on quantum computing, and he is also a Teaching By-Fellow at Churchill College, University of Cambridge. His research interests lie in quantum information, quantum computing, network information theory, machine learning and data science.

James R. Hopgood received the M.A., M.Eng. degree in Electrical and Information Sciences in 1997 and a Ph.D. in July 2001 in Statistical Signal Processing, part of Information Engineering, both from the University of Cambridge, England. He was then a Post-Doctoral Research Associate for the year after his Ph.D within the same group, at which point he became a Research Fellow at Queens College continuing his research in the Signal Processing Laboratory in Cambridge. He joined the University of Edinburgh in April 2004, and is currently a Senior Lecturer in the Institute for Digital Communications, within the School of Engineering, at the University of Edinburgh, Scotland. His research specialisation include model-based Bayesian signal processing, speech and audio signal processing in adverse acoustic environments, including blind dereverberation and multi-target acoustic source localisation and tracking, single channel signal separation, distant speech recognition, audio-visual fusion, medical imaging, blind image deconvolution, and general statistical signal and image processing. Since September 2011, he is Editor-in-Chief for the IET Journal of Signal Processing.

Professor Bernard (Bernie) Mulgrew (FIEEE, FREng, FRSE, FIET) received his B.Sc. degree in 1979 from Queen's University Belfast. After graduation, he worked for 4 years as a Development Engineer in the Radar Systems Department at Ferranti, Edinburgh. From 1983-1986 he was a research associate in the Department of Electrical Engineering at the University of Edinburgh. He was appointed to lectureship in 1986, received his Ph.D. in 1987, promoted to senior lecturer in 1994 and became a reader in 1996. The University of Edinburgh appointed him to a Personal Chair in October 1999 (Professor of Signals and Systems). His research interests are in adaptive signal processing and estimation theory and in their application to radar and sensor systems. Prof. Mulgrew is a co-author of three books on signal processing.